Europ. J. Combinatorics (1988) **9**, 575–579

Sur les Graphes Admettant le Nombre Maximum de Sous-Graphes à Trois Sommets et Deux Arêtes, et les Paires d'Ordres Totaux qui Maximisent |Rho − Tau|

C. LE CONTE DE POLY BARBUT

We show that the non-oriented graphs with a maximum number of 3-tuples of vertices with two edges in their subgraphs are the balanced bipartite complete graphs. This settles a conjecture formulated by G. Kreweras and allows to prove B. Monjardet's conjecture on the maximum difference between the tau and rho correlation coefficients on linear orders.

1. INTRODUCTION ET RAPPELS

Soucieux l'un et l'autre de comparer deux ordres totaux O_1, O_2, relatifs à un même ensemble E de n éléments, Kendall et Spearman [2] ont proposé deux techniques différentes.

Kendall considère les $N = n(n-1)/2$ paires d'éléments de E. Comptant le nombre a des accords et celui $d = N - a$ des désaccords entre O_1 et O_2 relativement au classement des éléments de ces paires, il définit son fameux coefficient de corrélation *tau* par la formule:

$$\tau(O_1, O_2) = (a - d)/N = 1 - 4\, d/n(n-1). \tag{1}$$

Spearman pour sa part considère les rangs $r_1(i)$ et $r_2(i)$ qu'assignent au même élément i les ordres O_1 et O_2, puis leur associe une distance d_s telle que

$$d_s^2(O_1, O_2) = \sum_{i \in E} [r_1(i) - r_2(i)]^2.$$

Lorsque O_1 et O_2 sont inverses l'un de l'autre, d_s^2 prend sa valeur maximum $M = (n^3 - n)/6$. Spearman définit alors son non moins fameux coefficient de corrélation *rho* par la formule

$$\varrho(O_1, O_2) = 1 - d_s^2(O_1, O_2)/M = 1 - 6d_s^2/(n^3 - n). \tag{2}$$

Il est clair que l'on a $0 \leqslant |\tau|, |\varrho| \leqslant 1$. Chacun des coefficients peut d'ailleurs s'interpréter comme un cosinus dans un espace convenable, ce qui éclaire leur nature, mais comme il faut faire appel à deux espaces *différents*, la comparaison ne s'en trouve pas facilitée pour autant.

Monjardet [3] propose de procéder à cette comparaison en utilisant l'ordre $O_1 \cap O_2$, partiel cette fois (dès que $O_1 \neq O_2$), d'associer à cet ordre son graphe de comparabilité noté $G = G(O_1, O_2)$ et d'en étudier les triplets de sommets.

Remarquons que, pour tout graphe simple (= non-orienté, sans boucle), les triplets de sommets se répartissent en quatre types de configuration (sous-graphes), selon le nombre i d'arêtes reliant les sommets du triplet:

$i = 0$ — Type D_1 | $i = 1$ — Type A_2 | $i = 2$ — Type D_2 ou configuration en V | $i = 3$ — Type A_1

Les configurations A_i et D_i sont arêtes complémentaires, $i \in \{1, 2\}$.

0195-6698/88/060575+05 $02.00/0

On peut dès lors attacher à tout graphe quatre entiers naturels d_1, a_1, d_2, a_2 qui donnent respectivement les nombres de configurations de type D_1, A_1, D_2, A_2 qui y figurent.

Il existe entre les deux coefficients ϱ et τ la relation suivante calculéee dans $G(O_1, O_2)$:

$$\varrho(O_1, O_2) - \tau(O_1, O_2) = 4(d_2 - a_2)/(n^3 - n) \qquad \text{(B. Monjardet, 1984)} \qquad (3)$$

Cherchant alors les paires $\{O_1, O_2\}$ qui maximisent $|\varrho - \tau|$, B. Monjardet formule les deux conjectures (C) et (K) suivantes, en remarquant que si (K) est vraie (C) l'est également (cf. [3]).

(C) $|\varrho - \tau|$ *est maximum pour les ordres totaux* O_1, O_2 *tels que* $G(O_1, O_2)$ *soit ou bien un graphe biparti complet équilibré, ou bien le complémentaire d'un tel graphe.*

Ici 'équilibré' signifie que les deux classes de sommets ont même cardinalité à un point près.

(K) *Pour n fixé les ordres partiels* P *à n éléments tels que le graphe de comparabilité* G(P) *admette le plus grand nombre de configurations en* V *sont les ordres* P *dont* G(P) *est biparti complet équilibré.*

G. Kreweras étend cette conjecture au cas des graphes non-orientés quelconques, soit:

(K′) *Pour n fixé les graphes non-orientés à n sommets qui admettent le plus grand nombre de configurations en* V *sont les graphes bipartis complets equilibrés.*

L'objet de ce travail est de démontrer la conjecture (K′) (Théorème 2.4).

2. Preuve de la Conjecture K′

(2.1) Notations. G = (V, E) étant un graphe sommets et arêtes simple, avec Card V = n:

(1) Une C E V (configuration en V) $x\,\bar{y}\,z$ est définie par:

$$x\,\bar{y}\,z \Leftrightarrow ((x, y) \in E, (y, z) \in E, x \neq z, (x, z) \notin E))$$

y est la *pointe* de la C E V ($x\,\bar{y}\,z$). (Une C E V est un sous graphe à trois sommets et deux arêtes.)

(2) Le *poids* de G, $p(G)$ est le nombre de C E V de G (paramètre d_2).

(3) Le *poids du sommet* x de G, $p(xG)$ est le nombre de C E V dans lesquelles x entre en tant que sommet, à la pointe ou non.

(4) (G/x) est le graphe obtenu en enlevant à G le sommet x et les arêtes qui lui sont adjacentes.

(5) Un graphe ayant un nombre maximum de C E V, pour n donné, est dit *maximal.*

(2.2) Remarques Préliminaires.

(1) $p(G) = \frac{1}{3}\Sigma_{x\in V}\, p(xG)$; la division par trois s'explique par le fait qu'une même C E V intervient trois fois sous le signe Σ.

(2) $p(G) = p(xG) + p(G/x)$; car une C E V appartient à (G/x) si et seulement si elle ne contient pas x.

(2.3) Proposition. *Dans un graphe maximal, deux sommets de poids différents sont nécessairement adjacents.*

PREUVE. Soit $G = (V, E)$ un graphe maximal. Soit encore x et y deux sommets tels que $p(xG) < p(yG)$; du fait que l'on a $1 \leqslant p(yG)$, le sommet y n'est pas isolé et admet le voisinage $\{z_1, \ldots, z_q\} \neq \varnothing$.

Supposons x et y non-adjacents, donc $x \notin \{z_1, \ldots, z_q\}$ et désignons par $\{t_1, \ldots, t_r\}$ les points situés à distance 2 de y. Il n'est pas exclu que l'on ait $x \in \{t_1, \ldots, t_r\}$.

Désignons par $G' = (V, E')$ le graphe obtenu en ajoutant à (G/x), x et les arêtes reliant x à tous les voisins de y. Par construction $(G'/x) = (G/x)$, donc $p(G'/x) = p(G/x)$ et $p(G') = p(xG') + p(G'/x) = p(xG') + p(G/x)$.

Nous définissons une injection naturelle de Y, l'ensemble des C E V de G contenant y, dans X′, l'ensemble des C E V de G′ contenant x:

(1) y est une pointe. Si $(z_i \, \bar{y} \, z_j) \in Y$ alors $(z_i \, \bar{x} \, z_j) \in X'$.

(2) y n'est pas une pointe.

(2a) $(y \, \bar{z}_i \, t_\lambda) \in Y$ et $x \neq t_\lambda$, alors $(x \, \bar{z}_i \, t_\lambda) \in X'$.

(2b) $(y \, \bar{z}_i \, x) \in Y$, alors $(y \, \bar{z}_i \, x) \in X'$.

L'existence de cette injection implique

$$p(xG') \geqslant p(yG) > p(xG)$$

donc

$$p(G') > p(G)$$

ce qui est contraire à la maximalité de G. □

COROLLAIRE. *Dans un graphe maximal, deux sommets non-adjacents sont nécessairement de même poids.*

(2.4) PROPOSITION. *Dans un graphe maximal, deux sommets non-adjacents ont les mêmes voisins (eux-mêmes mis à part).*

PREUVE. Soit x et y deux sommets non-adjacents d'un graphe G maximal, D'après ce qui précède on a $p(xG) = p(yG)$. Supposons que les voisinages de x et de y ne coincident pas. Sans perte de généralité on conviendra que le voisin z de y n'est pas adjacent à x.

La construction de $G' = (V, E')$ faite dans la preuve de la Proposition (2.3) prouve que l'on a

$$p(xG') \geqslant p(yG) = p(xG).$$

En outre, l'existence dans X′ de la C E V $(x \, \bar{z} \, y)$ montre que l'injection naturelle de Y dans X′ n'est pas une surjection. On en déduit $p(G') > p(G)$ ce qui contredit la maximalité de G.

COROLLAIRE. *Le complémentaire d'un graphe maximal est le graphe d'une relation d'équivalence.*

PREUVE. Soit $G = (V, E)$ un graphe maximal, $\bar{G} = (V, \bar{E})$ son complémentaire. Ecrivons $x \mathrel{R} y$ pour $(x, y) \in E$. Par définition $\rceil(x \mathrel{R} y) \Leftrightarrow x \mathrel{\bar{R}} y$ exprime $(x, y) \in \bar{E}$. La proposition précédente s'écrit

$$(\forall\, x, y, z \in V)\; x \mathrel{\bar{R}} y \Rightarrow [(x \mathrel{R} z \;\&\; z \neq y) \Rightarrow z \mathrel{R} y].$$

Le corollaire annoncé s'obtient immédiatement en contraposant cette implication.

CONSÉQUENCE. le complémentaire d'un maximal ne peut être qu'une union disjointe de graphes complets $K_{n_1}, K_{n_2}, \ldots, K_{n_m}$. Un maximal ne peut donc être qu'un graphe m-parti complet $K_{n_1,n_2,\ldots,n_m}$ relatif à une partition de V en classes $C_1, \ldots, C_m$, de cardinaux $n_1, n_2, \ldots, n_m$. Sans perte de généralité, nous supposerons par la suite $n_1 \leqslant n_2 \leqslant \ldots \leqslant n_m$.

(2.5) THÉORÈME. *Une condition nécessaire et suffisante pour qu'un graphe soit maximal est qu'il soit biparti complet équilibré.*

PREUVE. Soit G un graphe maximal; il est m-parti complet. Supposons $m > 3$ et considérons le graphe $\hat{G}$ obtenu à partir de G en supprimant les $n_1 \times n_2$ arêtes de G reliant C_1 à C_2. En passant de G à $\hat{G}$, cette suppression d'arêtes fait perdre

$$n_1\binom{n_2}{2} + n_2\binom{n_1}{2} = n_1 n_2\left(\frac{n_1 + n_2}{2} - 1\right) \text{C E V entre } C_1 \text{ et } C_2;$$

en revanche, elle en fait gagner $n_1 \times n_2 \times n_i$ pour tout triplet (C_1, C_2, C_i), $i \geqslant 3$. Or $n_1, n_2 \leqslant n_i$ entraîne $(n_1 + n_2)/2 - 1 < n_i$. On a donc $p(G) < p(\hat{G})$, ce qui est incompatible avec la maximalité de G.

Il en résulte qu'un graphe maximal a au plus deux classes. Il ne peut pas n'en avoir qu'une, car il serait dépourvu d'arêtes. Il en a donc exactement deux. Un tel graphe K_{n_1,n_2} admet $n_1 n_2((n_1 + n_2)/2 - 1)$ C E V. Pour $n = n_1 + n_2$ fixé, le maximum est obtenu en donnant à $n_2 - n_1$ la plus petite valeur possible, c'est-à-dire 0 ou 1 selon que n est pair ou impair: le graphe maximal est donc *équilibré* au sens défini dans l'introduction. □

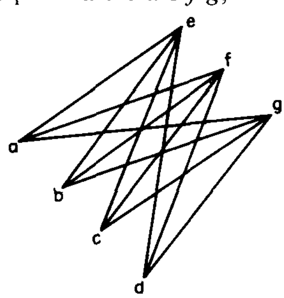
(2.6) REMARQUE.
Les graphes bipartis complets équilibrés sont aussi les graphes sans triangle admettant pour un nombre n de sommets fixé, le plus grand nombre d'arêtes (Bollobas [1], ch. III).

3. RETOUR A LA FORMULE $\varrho(O_1, O_2) - \tau(O_1, O_2) = 4(d_2 - a_2)/(n^3 - n)$

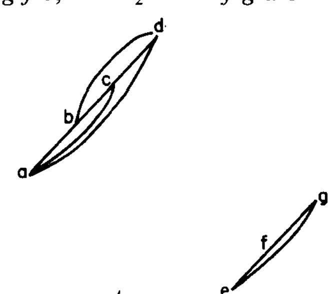
Un graphe G biparti complet équilibré ne contient aucune configuration du type A_2. G est alors graphe d'intersection de deux ordres avec d_2 maximum et a_2 nul; et $|\varrho - \tau|$ est maximum pour ces ordres. $\bar{G}$, le graphe complémentaire, union disjointe de deux graphes complets, équilibrés, est également graphe d'intersection de deux ordres. $\bar{G}$ contient un nombre maximum de configurations de type A_2 (ce nombre est égal au nombre d_2 de C E V de G) et ne contient aucune C E V. $|\varrho - \tau|$ est encore maximum pour les ordres dont G est graphe d'intersection.

EXEMPLES. Pour $n = 7$, $|\varrho - \tau|$ est maximum en particulier pour les deux couples d'ordres (O_1, O_2) et (O_1, O_2') dont les graphes $G(O_1, O_2)$ et $G(O_1, O_2')$ sont complémentaires l'un de l'autre—l'un étant biparti complet équilibré et l'autre union de deux graphes complets équilibrés:

$$O_1 = a\,b\,c\,d\,e\,f\,g, \qquad O_2 = d\,c\,b\,a\,g\,f\,e, \qquad O_2' = e\,f\,g\,a\,b\,c\,d.$$



G(O_1, O_2)



G(O_1, O_2')

O_1 est lu en projection horizontale; O_2 et O_2' en projection verticale.

Remerciements

Je remercie le Professeur André Lentin dont les remarques m'ont permis d'améliorer la rédaction de cette note.

References

1. B. Bollobas, *Extremal Graph Theory*, Academic Press, 1978.
2. M. G. Kendall, *Rank Correlation Methods*, Charles Griffin, 4e ed., 1970.
3. B. Monjardet et C. le Conte de Poly Barbut, Valeurs extrémales de la différence de deux coefficients de corrélation de rangs ϱ et τ, *C.R. Acad. Sci. Paris*, **303**, I, 10 (1986), 483–486.

Received 3 February 1987

C. Le Conte de Poly Barbut
Centre d'Analyse et de Mathématique Sociales,
Ecole des Hautes Etudes en Sciences Sociales,
54 bd Raspail, 75006 Paris, France